THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 4, Issue of January 28, pp. 2895-2901, 1994
Printed in U.S.A.
© 1994 by The American Society for Biochemistry and Molecular Biology, Inc.

An *Escherichia coli* Protein Consisting of a Domain Homologous to FK506-binding Proteins (FKBP) and a New Metal Binding Motif*

(Received for publication, July 15, 1993, and in revised form, September, 14, 1993)

Christoph Wülfing, Josefa Lombardero‡, and Andreas Plückthun§

From the Protein Engineering Group, Max-Planck-Institut für Biochemie, Am Klopferspitz, D-8033 Martinsried, Federal Republic of Germany

Initially detected as a persistent contaminant in immobilized metal affinity chromatography of recombinant proteins in *Escherichia coli*, a 196-amino acid protein was isolated, cloned, overexpressed, and characterized. It consists of two domains, of which the first (146 amino acids) shows some homology to the FK506-binding proteins. The second domain (50 amino acids) is extremely rich in potentially metal-binding amino acids, such as histidine, cysteine, and acidic amino acids. The protein binds Ni^{2+} and Zn^{2+} tightly with 1:1 stoichiometry, Cu^{2+} and Co^{2+} with lower affinity, and Mn^{2+}, Fe^{2+}, Fe^{3+}, Mg^{2+}, and Ca^{2+} hardly at all.

Metals can be bound by a large number of diverse proteins. They may have an exclusively structural role as *e.g.* in the zinc-finger domains of several transcription factors, or they may be involved in the catalytic function of the protein. Most metals can serve structural functions, but redox-active metals are less likely to do so, since the change in redox state can change the preferences for coordination geometry. Catalytic functions of metals include redox and Lewis acid chemistry, both of which are more likely to be carried out by transition group metals due to their ease of change of redox state and their polarizing capability.

During extensive work on the purification of several heterologous recombinant proteins from *E. coli* by immobilized metal affinity chromatography (IMAC)[1] (Porath, 1987; Arnold, 1991; Lindner *et al.*, 1992), a host-encoded protein was observed as the major contaminant and could be isolated and purified to homogeneity. In view of its metal binding ability expected from its binding to the IMAC column, and taking into account the general importance of transition metal binding proteins in *E. coli* (Beveridge, 1989), this protein was cloned. We found that it consists of two domains, an NH_2-terminal domain of about 150 amino acids and a COOH-terminal metal binding motif of about 50 amino acids.

During the characterization of this protein a weak, but significant homology of the NH_2-terminal domain to FK506-binding proteins (Harding *et al.*, 1986; Siekierka *et al.*, 1989) was detected, a class of proteins involved in immunosuppression. Two classes of such proteins have been described, the other being cyclophilins (Fischer *et al.*, 1984) which bind the immunosuppressant cyclosporin A. The two classes are related neither by sequence homology nor folding topology, nor is there any conspicuous structural homology between the immunosuppressants FK506 and cyclosporin A. Both classes do show proline *cis-trans*-isomerase (rotamase) activity and can accelerate the folding of some proteins *in vitro*. The biological significance of this *in vitro* reaction for immunosuppression remains a mystery, however. Rather, the complex of immunosuppressant and immunophilins seems to react with phosphatases (Liu *et al.*, 1991). Both classes include bacterial homologs (Hayano *et al.*, 1991; Liu and Walsh, 1990; Perry *et al.*, 1988) that are not all able to bind the respective immunosuppressants.

We describe here the isolation, cloning, sequence, overexpression, and native purification of the protein. It was characterized by CD spectroscopy, and its metal binding properties were investigated.

MATERIALS AND METHODS

DNA Manipulations—DNA manipulations were performed according to standard protocols (Sambrook *et al.*, 1989). The PCR oligonucleotides derived from the peptide sequences were ACT GAG AAT TCT AYC ARG TN(C/A) GNA CNG ARG A and CGG ACG TCG ACY TCY TCY TCN GTN GCY TC. The product of the first PCR with the degenerate oligonucleotides and *Xba*I-cleaved genomic DNA was gel-purified and cloned into the sequencing vector pSL301 (Invitrogen Corp., San Diego, CA) after cleavage with the restriction endonucleases *Eco*RI and *Xba*I.

Inverse PCR (Silver, 1991) was performed in two parts after determining suitable restriction enzymes by Southern blotting (Gebeyehu *et al.*, 1987). Using the restriction endonuclease *Afl*III, which cuts within the already known part of the gene for the preparation of the inverse PCR template, two independent fragments, a 800-base pair fragment for the NH_2 terminus and a 3.5-kilobase pair fragment for the COOH terminus, were amplified. The PCR oligonucleotides for the inverse PCR amplification of the NH_2 terminus were ACT GAG AAT TCG GCG CGA ACG ACG CTT and ACT GAT CTA GAT CAC CGG AGA CTC ATC AA, and the ones for the COOH terminus were ACT GAG AAT TCT GAA ATT CAA CGT TGA AGT and ACT GAT CTA GAC CCA TAA ATA CGT CTT TAG G. The inverse PCR for the NH_2 terminus was performed following Ponce and Micol (1992), and the reaction product was cloned as above. For the COOH terminus the 3.5-kilobase pair DNA fragment was amplified following Kainz *et al.* (1992). After cleavage with *Nru*I and *Nsi*I and fill-in of the *Nsi*I ends, the reaction product was cloned blunt-ended into pSL301.

In order to confirm the WHP sequence, the gene was PCR-amplified three times directly from genomic DNA using the inward oligonucleotides ACT GAT CTA GAG GCT GAA GAA ACG CCA and ACT GAG AAT TCC GCT ACA ATC TGC G. After cloning via *Eco*RI and *Xba*I in pSL301, the three PCR products were sequenced.

In order to express WHP with sufficient yield, the gene was ligated into the *Sph*I ATG-cloning site of pUHE25-2.[2] For this purpose, the gene of WHP was cut out of the sequencing vector with *Eco*RV and *Xba*I and blunt-ended. The expression vector was termed pWHPexpr.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) Z21496.

‡ Present address: Center of Immunology, P. O. Box 16040, Havana 11600, Cuba.

§ To whom correspondence should be addressed (at his new address): Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland. Fax: 41-1-257-5712.

[1] The abbreviations used are: IMAC, immobilized metal affinity chromatography; PCR, polymerase chain reaction; iPCR, inverse PCR; MOPS, 4-morpholinepropanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; AAS, atomic absorption spectroscopy; ORF, open reading frame; FKBP, FK506-binding proteins.

[2] H. Bujard, personal communication.

Protein Purification—A 2-liter culture of *E. coli* strain RB791 (Brent and Ptashne, 1981) harboring plasmid pWHPexpr was grown to OD_{550} = 0.5, the expression of WHP was induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside, and the cells were grown for another 3 h at 37 °C. After cell rupture in a French pressure cell, the soluble fraction was loaded onto a Ni^{2+}-NTA column (25 ml), equilibrated with 100 mM Tris, pH 8.0, 1 M NaCl. After washing the column with 500 ml of equilibration buffer containing 8 mM imidazole, WHP was eluted with a 500-ml gradient from 8 to 100 mM imidazole. The column was run at a flow rate of 60 ml/h. The pooled WHP fractions were concentrated by ultrafiltration, and the buffer was changed on a PD10 column (Pharmacia) to 10 mM MOPS, pH 7.0. The sample was then loaded onto a DEAE-Sepharose column (66 ml) (Pharmacia), washed with 100 ml of 10 mM MOPS, pH 7.0, and the column was run with a gradient from 0 to 1 M NaCl. WHP eluted at about 0.5 M NaCl. The column was run at a flow rate of 35 ml/h.

Alternatively, a hydroxylapatite column was used as the first chromatography step. 500 ml of *E. coli* cells were grown as above. After centrifugation the bacterial cell pellet was resuspended in 1 mM NaCl. After cell rupture as above, the soluble fraction was loaded onto a 125-ml hydroxylapatite column (Bio-Rad). The column was washed, first with 200 ml of 1 M NaCl, then with 200 ml of 10 mM sodium phosphate, pH 6.8. WHP was eluted with a gradient of 600 ml sodium phosphate (10–300 mM), pH 6.8. As a second purification step a DEAE-Sepharose column was used as above.

Rotamase Activity Assay—The assay was performed according to Liu *et al.* (1990). WHP purified by IMAC was assayed at a final protein concentration of 0.2 μM without a detectable rate acceleration in the isomerization reaction. As a positive control, *E. coli* periplasmic peptidyl proline *cis-trans*-isomerase gave a strong acceleration in the assay already at 50-fold lower molar concentrations.

Computer Methods—Data base homology searches were performed using the UWGCG (University of Wisconsin Genetics Computer Group) and PIR (Protein Identification Resource) program packages. The FASTA algorithm (Pearson and Lipman, 1988) was used for simple sequence homology searches and PROFILESEARCH (Gribskov *et al.*, 1989) to search for homology to a sequence profile, generated with PILEUP. This same algorithm was used to find sequences compatible with the three-dimensional profile, determined for the FK506-binding protein structure (van Duyne *et al.*, 1991) by the method of Eisenberg and co-workers (Lüthy *et al.*, 1992). The CLUSTAL (Higgins *et al.*, 1992) program was used for clustered sequence alignment. The isoelectric point was calculated using the UWGCG program package.

CD Spectroscopy—Samples were dialyzed against 5 mM sodium phosphate buffer, pH 7.5, before recording the spectrum. The concentration of WHP was 0.83 mg/ml. The spectrum was recorded between 192 and 250 nm. An Auto Dicrograph Mark IV Spectrometer from Jobin Yvon Division Instruments SA was used. The data were processed using the OMA software from Instruments SA (Provencher, 1982), including the CONTIN program to calculate secondary structure contents.

Metal Binding Studies—Protein concentrations were determined using OD_{280} with an extinction coefficient of $\epsilon = 5.6$ (mM cm)$^{-1}$ (Gill and von Hippel, 1989). For each metal binding assay, 1 mg of native IMAC-purified protein was dialyzed overnight against 1 mM EDTA, 20 mM Tris, 10 mM citrate, 65 mM NaCl, 250 mM KCl, pH 7.6. Subsequently, metal ions were added at a final concentration of 2 mM, and the metal was allowed to bind overnight. Excess metal was removed by gel filtration on a Sephadex G-25 column (50 cm × 1.4 cm). The column was equilibrated, washed, and eluted with 300 mM NaCl, 20 mM Tris, pH 7.5. The Fe^{2+} samples were allowed to equilibrate under rigorous exclusion of air in the presence of 20 mM ascorbate to prevent oxidation to Fe^{3+}. In one Fe^{2+} sample phenanthroline was present at a concentration of 10 mM to prevent precipitation of iron hydroxides. Finally, the metal content after removal of unbound metal ions by gel filtration was determined by atomic absorption spectroscopy on a 1100B Spectrometer from Perkin Elmer. As controls, samples taken before and after incubation with EDTA were measured as well. Absence of proteolytic degradation was confirmed by running an aliquot of the AAS samples on an SDS-PAGE gel.

RESULTS

Cloning of the Gene of WHP—In experiments involving purification of recombinant proteins with a his_5 tail on IMAC columns (Lindner *et al.*, 1992), a persistent contaminant, termed WHP, of an apparent molecular mass of 27 kDa, as judged by SDS-PAGE, was noted. WHP bound to the IMAC column in the presence of 6 M guanidine HCl. Since this treatment disrupts the structure of essentially all proteins, an affinity to the column would require Ni^{2+}-binding residues clustered not only in the structure, but also in the sequence. This fact seemed to be unusual enough to investigate the protein further.

The protein was subjected to NH_2-terminal protein sequencing and amino acid analysis. A peptide from a tryptic digest was also isolated using an anhydro-trypsin column (Ishii *et al.*, 1983), which usually results in the isolation of the COOH-terminal peptide, and it was sequenced as well.

For the PCR cloning of the gene of WHP from genomic DNA, two sets of degenerate oligonucleotides were used. The first was derived from the amino acid sequence of the NH_2 terminus of WHP, the second from the sequence of the major tryptic peptide in the run-through of the anhydro-trypsin affinity column (Ishii *et al.*, 1983), assuming that this peptide was the COOH-terminal one. A PCR fragment of only 450 base pairs was obtained, cloned, and sequenced, but its length indicated that the complete WHP could not have been encoded on this fragment, and a considerable part of the COOH terminus of the gene was still missing.

Southern blotting using the cloned PCR fragment as a probe was then carried out to determine which restriction enzymes give rise to fragments suitable for inverse PCR (iPCR) (Silver, 1991) which might contain the whole gene of WHP. Attempts to clone fragments from iPCR longer than 1200 base pairs failed. Only clones with large deletions could be isolated, indicating that a gene upstream or downstream of the WHP gene is sensitive to overexpression in the homologous system. To avoid this problem, the gene was assembled from parts using inverse PCR of restriction digests of genomic DNA that placed the respective termini of the WHP gene on different restriction fragments. The NH_2-terminal part of the gene could be cloned without problems; for the COOH terminus the iPCR product had to be shortened by restriction digestion to permit cloning. The gene sensitive to overexpression is thus located downstream of the WHP gene.

The sequence information obtained by iPCR was used to amplify and clone the whole WHP gene directly from the *E. coli* genome using inward primers. The final sequence of the coding region was determined from three independent inward PCRs and the inverse PCR reaction.

The NH_2-terminal part of the gene includes a putative translation initiation region and a possible promoter sequence (Fig. 1) with significant homology to *E. coli* σ^{70} promoters. The amino acid composition derived from translation of the gene corresponds to the experimentally determined one (data not shown), and the sequences of the two peptides, determined by Edman degradation (shown *underlined* in Fig. 1), were found in the DNA sequence. Comparing DNA and protein sequences, WHP does not have a signal peptide, it thus is a cytoplasmic protein. The existence of a transcription terminator consensus sequence right behind the stop codon (Fig. 1) shows that the complete gene has been cloned and that the gene is a separate transcription unit. An apparent molecular mass of 26 kDa on SDS-PAGE corresponds well with an acidic 196-amino acid protein (calculated isoelectric point = 4.85) with a calculated molecular mass of 22 kD.

The amino acid sequence was compared with the Swissprot data base. One putative 17-kDa translation product of an *E. coli* open reading frame (accession number P22563) and another one from a *Pseudomonas fluorescens* ORF (accession number P21863) are 30% identical to the first 150 amino acids of WHP (see below). This part of WHP is exactly encoded by the gene fragment, which was cloned after PCR amplification using the oligonucleotides derived from the NH_2-terminal and the presumed COOH-terminal peptide. The 450-base pair frag-

```
         .         .         .         .         .         .  60
AAAATAATATTGAGATTGTTGAATGTGTTAAGTGCGGACATCAGATGCGAGAAGCAGACA

         .         .         .         .         .         . 120
AAGAAGCCCGCGATCACGTTCGCAAAGATGAGCAAGTGATCGGGTTTTTATCCGGACTAG

         .         .         . -35     .         .     -10. 180
CGATATGCGCCGTGTTTTTTTAAGCTAGTGAGTACACGGCTGCAGGAATTCCGCTACAAT

         .         .      rbs.         .         .         . 240
CTGCGCCACTATTCTTCCCATGCTCAGGAGATATCATGAAAGTAGCAAAAGACCTGGTGG
                                      M  K  V  A  K  D  L  V  V

         .         .         .         .         .         . 300
TCAGCCTGGCCTATCAGGTACGTACAGAAGACGGTGTGTTGGTTGATGAGTCTCCGGTGA
 S  L  A  Y  Q  V  R  T  E  D  G  V  L  V  D  E  S  P  V  S

         .         .         .         .         .         . 360
GTGCGCCGCTGGACTACCTGCATGGTCACGGTTCCCTGATCTCTGGCCTGGAAACGGCGC
 A  P  L  D  Y  L  H  G  H  G  S  L  I  S  G  L  E  T  A  L

         .         .         .         .         .         . 420
TGGAAGGTCATGAAGTTGGCGACAAATTTGATGTCGCTGTTGGCGCGAACGACGCTTACG
 E  G  H  E  V  G  D  K  F  D  V  A  V  G  A  N  D  A  Y  G

         .         .         .         .         .         . 480
GTCAGTACGACGAAAACCTGGTGCAACGTGTTCCTAAAGACGTATTTATGGGCGTTGATG
 Q  Y  D  E  N  L  V  Q  R  V  P  K  D  V  F  M  G  V  D  E

         .         .         .         .         .         . 540
AACTGCAGGTAGGTATGCGTTTCCTGGCTGAAACCGACCAGGGTCCGGTACCGGTTGAAA
 L  Q  V  G  M  R  F  L  A  E  T  D  Q  G  P  V  P  V  E  I

         .         .         .         .         .         . 600
TCACTGCGGTTGAAGACGATCACGTCGTGGTTGATGGTAACCACATGCTGGCCGGTCAGA
 T  A  V  E  D  D  H  V  V  V  D  G  N  H  M  L  A  G  Q  N

         .         .         .         .         .         . 660
ACCTGAAATTCAACGTTGAAGTTGTGGCGATTCGCGAAGCGACTGAAGAAGAACTGGCTC
 L  K  F  N  V  E  V  V  A  I  R  E  A  T  E  E  E |L  A  H
                                                   | metal-binding
         .         .         .         .         .         . 720
ATGGTCACGTTCACGGCGCGCACGATCACCACCACGATCACGACCACGACGGTTGCTGCG
 G  H  V  H  G  A  H  D  H  H  H  D  H  D  H  D  G  C  C  G
domain
         .         .         .         .         .         . 780
GCGGTCATGGCCACGATCACGGTCATGAACACGGTGGCGAAGGCTGCTGTGGCGGTAAAG
 G  H  G  H  D  H  G  H  E  H  G  G  E  G  C  C  G  G  K  G

         .         .         .         .         .         . 840
GCAACGGCGGTTGCGGTTGCCACTAATACCGAAAAAGTGACAAAAAAGCGGGGAATCCCC
 N  G  G  C  G  C  H  *                 ----------->   <---

         .         .         .         .         . 890
GCTTTTTTTACGCCTCAATAATGTGGCGGTGGCGTTTCTTCAGCCTGCGA
--------
```

FIG. 1. **Nucleotide and amino acid sequence of WHP.** A putative ribosome binding site (Scherer *et al.*, 1980; Stormo *et al.*, 1982) is *underlined* and indicated (*rbs*). A possible promoter sequence (Harley and Reynolds, 1987) has been defined by homology (indicated by *−35* and *−10*), and the sequence of a putative terminator, found by computer search (Brendel and Trifonov, 1984), is indicated by *two arrows*. The beginning of the second domain is indicated. The peptide sequences determined by Edman degradation are *underlined*.

ment cloned after the first PCR obviously only encodes the major NH_2-terminal degradation product of WHP. The sequence of the NH_2-terminal degradation product, as defined by the two tryptic peptides, exactly coincides with the region of homology between WHP and the two ORFs, suggesting that the NH_2 terminus is a distinct domain.

The COOH-terminal part of WHP contains a cluster of potentially metal-chelating amino acids such as histidine, cysteine, and aspartic and glutamic acid, suggesting that one or several metal ions are bound there. Since WHP is a cytoplasmic protein the cysteines will be in the reduced form and can potentially take part in chelating metal ions. A direct internal repeat (Fig. 2*a*) within this domain can be noted. The sequence of the COOH terminus of WHP allows for the possibility that different metal ions are bound by different residues of WHP, according to their coordination requirements.

Overexpression and Purification of Recombinant WHP—To characterize WHP further, an overexpression system was constructed. Since WHP is a soluble cytoplasmic protein, a direct expression system using a strong repressible promotor seemed to be the obvious choice. The gene encoding WHP was cloned into the vector pUHE25-2.[2] The vector provides the very strong A1 phage promoter made repressible by the introduction of *lac* repressor binding sites. The gene was inserted right after the phage T5-derived ribosomal binding site II provided by the vector, resulting in the plasmid pWHPexpr. RB791 (Brent and Ptashne, 1981), a LacI-overproducing *E. coli* W3110 derivative, was used as expression strain.

The recombinant protein was expressed with a yield of 50 mg/liter culture. It was found to remain soluble, and no deleterious effect on the cells was noted. The protein was purified by IMAC on a Ni^{2+}-NTA column under native conditions as de-

a

```
147                                 L  A  H  G  H  V  H  G  A

156   H  D  H  H  H  D  H  D  H  D  G  C  C  G  G  H  G
      |  |  |     |  :  |        :  |  |  |  |  |     |
173   H  D  H  G  H  E  H  G  G  E  G  C  C  G  G  K  G

190   N  G  G  C  G  C  H
```

Amino acid composition	H	30.0%	complete domain
	H	32.4%	internal repeat
	C	11.8%	
	E+D	20.6%	

b

WHP

```
122                                 ...NHMLAGQNLKFNVEVVAIREA
147   TEEELAHGHVAPAHDHHHDHDHDGCCGGHGHDHGHEHGGEGCCGGKGNGGCGCH

1     MCTVCGCGTSAIEGHTHEVGDDGHGHHHHDGHHDHDHDHDHHRGDHEHDDHHHA
54    EDGSVHYSKGIAGVHVPGMSQERIIQ...
```

Hydrogenase operon ORF

FIG. 2. *a*, special features of the metal binding domain of WHP. A direct repeat of 17 amino acids (amino acids 156–172 and amino acids 173–189) is *highlighted*. Identical amino acids are labeled with *vertical bars* and similar ones with *colons*. The two repeats contain 67% potentially metal ion binding amino acids. *b*, sequences of the COOH terminus of WHP and the NH_2 terminus of the translation product of an ORF from the hydrogenase operon from *R. leguminosarum* are shown in comparison in the *center lines*. Adjacent sequences of the respective proteins are shown in the *outer lines*. Potentially metal-chelating amino acids are *highlighted*.

scribed under "Material and Methods." Under these conditions, however, other host proteins still slightly contaminate the WHP fractions after IMAC, and a DEAE ion exchange chromatography step was used to obtain pure protein (Fig. 3). It is with this protein that the metal binding reconstitution studies were performed. As an alternative, in order not to introduce Ni^{2+} during the purification, WHP was also purified to homogeneity by a combination of chromatography on a hydroxylapatite and a DEAE ion exchange column as described under "Material and Methods."

Metal Binding Assays—In order to determine whether there are metal ions bound by WHP after isolation from *E. coli,* the metal content of WHP solutions was assayed by AAS. For this purpose, samples were taken at the end of the native purification procedure involving IMAC: they contained Ni^{2+} and Zn^{2+}. Whereas Ni^{2+} could have, but not necessarily has been, introduced during Ni^{2+} IMAC purification by a metal exchange process, Zn^{2+} must have been bound to WHP already in the cell.

Similarly, samples were taken at the end of the purification procedure involving the hydroxylapatite column chromatography. In this case, Ni^{2+}, Zn^{2+}, and Ca^{2+}, but no Mn^{2+}, Co^{2+}, or Cu^{2+}, were found to be bound to WHP (data not shown). Using a similar argument as above Ca^{2+} may have been introduced by the hydroxylapatite column, whereas Zn^{2+} and Ni^{2+} should have been bound to WHP already *in vivo*.

To determine which metal ion *can* bind to WHP, the protein, purified using IMAC, was stripped of metals by dialysis against an EDTA containing buffer, physiological in pH and salt concentration, and then incubated with a large molar excess of the metal ion to be tested. Ni^{2+} and Zn^{2+} both could not be removed completely by dialysis against EDTA, whereas subsequent incubation with a large molar excess of other divalent metal ions lead to a complete removal of Ni^{2+} and Zn^{2+} in all cases assayed (Table I). This indicates that metal ions can most efficiently be replaced via a direct exchange process. The metal binding domain of WHP thus seems to have a very strong need for a bound metal ion, but does not seem to care very much about the type of metal.

In the metal reconstitution experiments, Ni^{2+} and Zn^{2+} were the metals being bound best, but Cu^{2+} and Co^{2+} also bound to WHP, albeit to a smaller extent (Table I). It is noteworthy that Ni^{2+} and Zn^{2+} bind to WHP with 1:1 stoichiometry, leading to the conclusion that only one metal binding site is present in WHP despite the fact that a motif is present in direct repeat in the WHP metal binding domain. Iron did not bind to WHP at all, no matter whether it was added as FeIII or as FeII, complexed weakly by citrate or strongly by phenanthroline in the reconstitution experiment. Oxidation of FeII in slightly alkaline solution was prevented by addition of ascorbate. Mg^{2+} and Ca^{2+} did not bind to WHP under these conditions.

Comparing the metal reconstitution data to the data obtained by assaying WHP solutions at the end of the respective purification procedures, it is found that the results generally coincide well. Both kinds of experiments support the idea that Ni^{2+} and Zn^{2+} are preferentially bound. Yet in contrast to the metal reconstitution experiments, some Ca^{2+} was found to be bound to WHP when purified using the hydroxylapatite column. The high Ca^{2+} concentrations on the column seem to force WHP to accept the less preferred metal ion as a ligand. This again suggests that it is most important for WHP to have a metal ion bound at all.

In conclusion, the metal binding domain of WHP seems to be able to bind divalent metal ions within a certain range of ionic radii. Copper is bound to a smaller extent probably due to its strict geometrical coordination requirements. Iron is not bound at all, perhaps partly because of its small effective concentration in neutral or alkaline solutions due to oxo-bridging, and the affinity of WHP appears to be insufficient to sequester this metal. As Mg^{2+} is smaller and Ca^{2+} is larger than the divalent transition metals assayed, these metals do not seem to fit very well into the WHP metal binding site. Although a large number of oxygen ligands, preferred by these metals, are present in this domain, these metals bind only under extreme conditions or not at all.

Homology Search—The NH_2 terminus of WHP and two homologous uncharacterized open reading frames, one in *E. coli*, one in *P. fluorescens*, have about 30% identity compared with each other (Fig. 4). Using these three sequences, a homology profile was constructed, and the Swissprot data base was searched with this profile. A significant homology to, and only to, the FKBP family showed up.

Conversely, the three-dimensional homology profile of the FK506-binding protein (van Duyne *et al.*, 1991) was calculated according to the method of Eisenberg and co-workers (Lüthy *et al.*, 1992). This type of profile contains no sequence information at all, but is a measure of the compatibility of a sequence with a given three-dimensional structure. Searching the Swissprot data base with this profile, both the hypothetical 16.1-kDa protein of *E. coli* and the 16.3-kDa hypothetical protein of *P. fluorescens* were found with 8.4 and 7.1 standard deviations, second only to the FKBP family. Next a cluster alignment of the WHP, the 16.1-kDa and the 16.3-kDa protein and the bovine FKBP sequence, was performed (Fig. 4). It appears as if there might be an insertion in the structure of FKBP (Figs. 4 and 5).

To obtain some initial experimental information about the protein structure, CD spectra were recorded (Fig. 6). They suggest a content of β-sheet of 51 ± 8% and 10 ± 4% α-helix, and this would be consistent with a chain fold similar to FKBP, albeit with the uncertainty of the contribution of the large

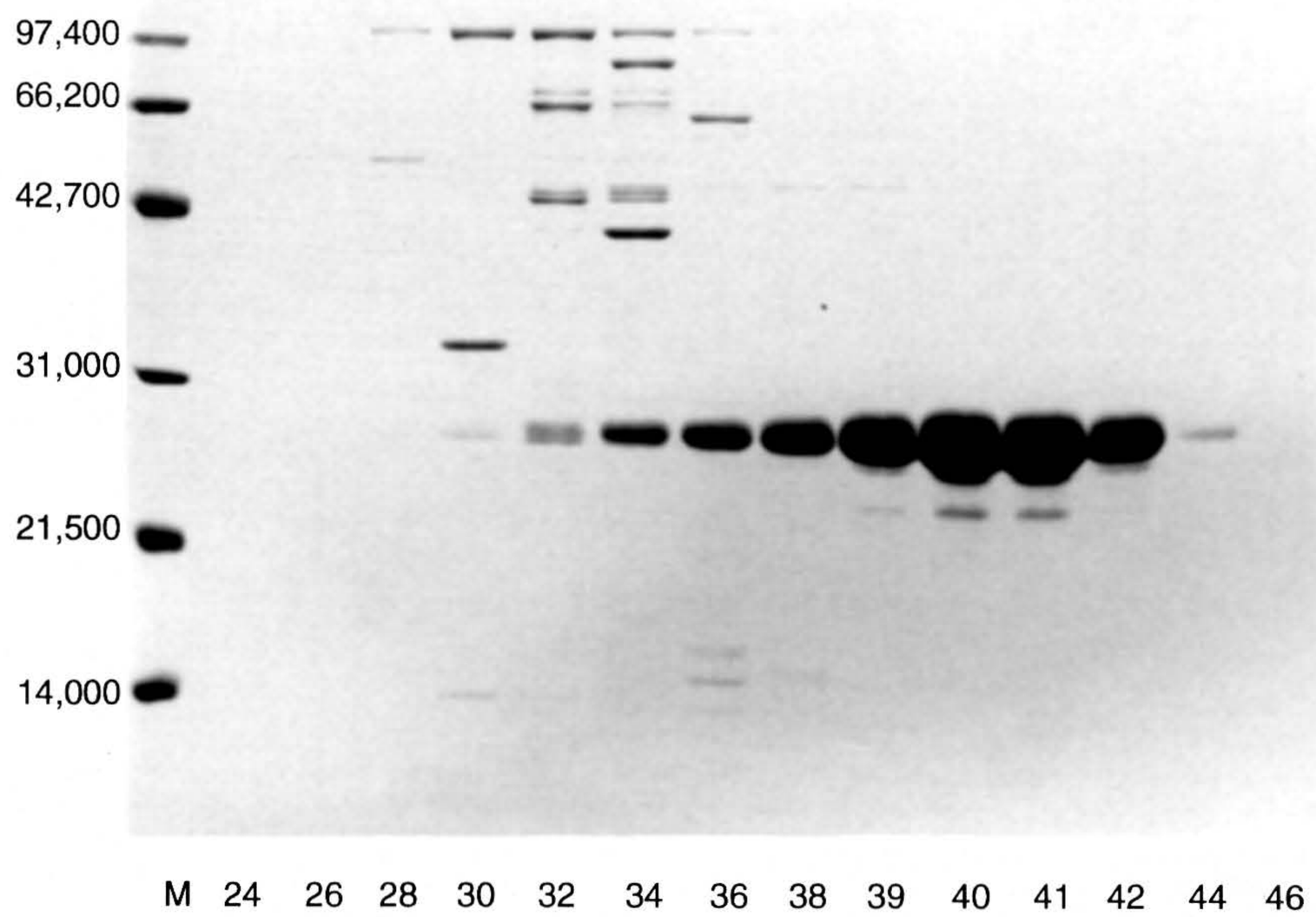

FIG. 3. **Purification of WHP.** Shown is an SDS-PAGE gel of different fractions of the DEAE ion exchange column, used as the last step in the native purification procedure. The fraction number from a salt gradient is given below each lane. *M* denotes the molecular weight marker.

TABLE I
Metal binding assays

WHP was stripped of its metal ions and was then reconstituted by incubation with a solution containing the metal ion specified in the first column. "EDTA" denotes that the metal was stripped from WHP without later metal ion addition, "none" denotes WHP receiving no treatment at all. The WHP samples treated according to the first column were assayed for their metal content by AAS, the type of metal ion being assayed is given in each column heading. Numbers give micromolar concentrations; ND, no determination. The numbers in the last column give the micromolar concentrations of WHP in the respective AAS samples.

Treatment	Concentration of metal analyzed in WHP sample								WHP concentration
	Mg^{2+}	Zn^{2+}	Cu^{2+}	Ni^{2+}	Co^{2+}	Fe^{2+}	Fe^{3+}	Ca^{2+}	
				μM					*μM*
Mg^{2+}	0	ND	ND	ND	ND	ND	ND	ND	52
Zn^{2+}	ND	32	0	0	0	ND	ND	ND	30
Cu^{2+}	ND	0	37	0	0	ND	ND	ND	57
Ni^{2+}	ND	0	0	61	0	ND	ND	ND	57
Co^{2+}	ND	0	0	0	15	ND	ND	ND	54
Fe^{2+}	ND	ND	ND	ND	ND	0	ND	ND	47
Fe^{3+}	ND	ND	ND	ND	ND	ND	0	ND	40
Ca^{2+}	ND	ND	ND	ND	ND	ND	ND	0	88
EDTA	0	4	0	19	0	0	0	0	130
None	0	21	0	59	0	ND	ND	0	215

insertion and the metal binding domain. In order to investigate the possibility that the predicted fold, similar to the FKBP family, allows for rotamase activity, a standard proline *cis-trans*-isomerase activity assay (Liu *et al.*, 1990) was performed. No activity was found for WHP. It should be kept in mind, however, that the rotamase activity may be specific for a different type of substrate, and this cannot be excluded from this experiment.

Searching the data base with the COOH-terminal fragment and different amino acid patterns thereof led to the 50 NH_2-terminal amino acids of an open reading frame from the hydrogenase operon of *Rhizobium leguminosarum* (EMBL accession number: X52974) which is also rich in acidic amino acids, cysteine, and histidine (Hidalgo *et al.*, 1992) (Fig. 2*b*). This ORF has three close homologs (PIR international data base accession numbers S23440, S15198, D38532) which only differ in the respective NH_2-terminal parts (data not shown), suggesting that the NH_2 termini are variable extensions of a common fold. Although no homology was found which was significant enough to provide a structural model for the COOH-terminal part of WHP, the analogy to the ORF from the hydrogenase operon suggests that such NH_2- or COOH-terminal clusters of potentially metal-binding amino acids may form a separate functional domain.

DISCUSSION

WHP presumably is a two-domain protein, evidenced by homology and the facile proteolytic degradation at the putative domain border (see above). The NH_2-terminal domain seems to contain an FKBP folding motif. This is suggested by three lines of evidence. First, sequence homology searches with either the NH_2-terminal domain of WHP, or with the translation products of the homologous ORFs from *E. coli* and *P. fluorescens*, or with a sequence profile of the three aligned proteins shows that WHP and the two homologous ORFs are related to the FKBP family. Second, searching the data base with a FKBP structural profile (Lüthy *et al.*, 1992), not containing any sequence information, retrieves the two ORFs, WHP not being in the data base, as the best structures directly after FKBPs at about 7 and 8.5 standard deviations. Third, taking a cluster alignment of WHP, the ORFs, and FKBP based on sequence homology, the 50 amino acids by which the NH_2-terminal domain of WHP and its homologs are longer than FKBP are placed in a large loop of the FKBP structure where they can be accommodated (Figs. 4 and 5).

The COOH-terminal metal binding motif of WHP is remarkable because it contains a cluster of almost 30 potentially metal-binding amino acids. Experimentally, the binding of

```
Similarity to WHP     *            +  +  * + + ++ ++*   + +   + *  +  * ++*  *+*
WHP                 --M--------K---VAKDLVVSLAYQVRTEDGVLVDESPVSA-PLDYLHGHGSLISGLE   46
P. fluorescens      --MTDQVLAEQR---IGQNTEVTLHFALRLENGDTVDSTFDKA-PATFKVGDGNLLPGFE   54
E. coli             --MSE------S---VQSNSAVLVHFTLKLDDGTTAESTRNNGKPALFRLGDASLSEGLE   49
FKBP                GVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRNKPFKFMLGKQEVIRGWE   60
FKBP structure       ββββββββ            βββββββββββ    βββββ                  αααα
Similarity to FKBP                          +*+   ****   **+     *  *  *   +  *+*

Similarity to WHP   + *   ***    +    *+*      +* + +  *      + * + ++       +
WHP                 TAMEGHEVGDKFDVAVGANDAYGQYDENLVQRVPKDVFMGVDELQVG-MRFLAETDQGPV  105
P. fluorescens      AALFGFKAGDKRTLQILPENAFGQPNPQNVQIIPRSQFQNMDLSE-GLLVIFNDAANTEL  113
E. coli             QHLLGLKVGDKTTFSLEPDAAFGVPSPDLIQYFSRREFMDAGEPEIGAIMLFTAMDGSEM  109
FKBP                EGVAQMSVGQRAKLTISPDYAYGATGH--------------------------------   87
FKBP structure      ααα        βββββββ
Similarity to FKBP   +    **++  + + ++ *+*

Similarity to WHP   *  +  +    + +* ** ***  + *++*++ +
WHP                 PVEITAVEDDHVVVDGNHMLAGQNLKFNVEVVAIREATEEE  147
P. fluorescens      PGVVKAFDDAQVTIDFNHPLAGKTLTFDVEIIDVKAL----  151
E. coli             PGVIREINGDSITVDFNHPLAGQTVHFDIEVLEIDPALEA-  150
FKBP                PGIIPPHA---------------TLVFDVELLKLE------  107
FKBP structure                                ββββββββββββ
Similarity to FKBP  ** +                      ** ****++ +
```

FIG. 4. **Sequence alignment of WHP, the homologous ORFs from *E. coli* (Swissprot accession number P22563) and *P. fluorescens* (Swissprot accession number P21863), and the human FK506-binding protein (Brookhaven data base entry 1FKF), as performed by CLUSTAL.** The line on top of the WHP sequence shows positions at which the first three proteins of the alignment are similar (+) or identical (*). For comparison, the amino acids being identical or similar between all four aligned proteins are marked in the last line denoted "*Similarity to FKBP*." A *star* signifies an amino acid that is identical between FKBP and at least two of the aligned proteins; a *plus* denotes 4 similar amino acids. The line called "*FKBP structure*" shows the amino acids being involved in elements of secondary structure as determined by x-ray crystallography. α denotes α-helix, and β denotes β-sheet as the respective secondary structure element.

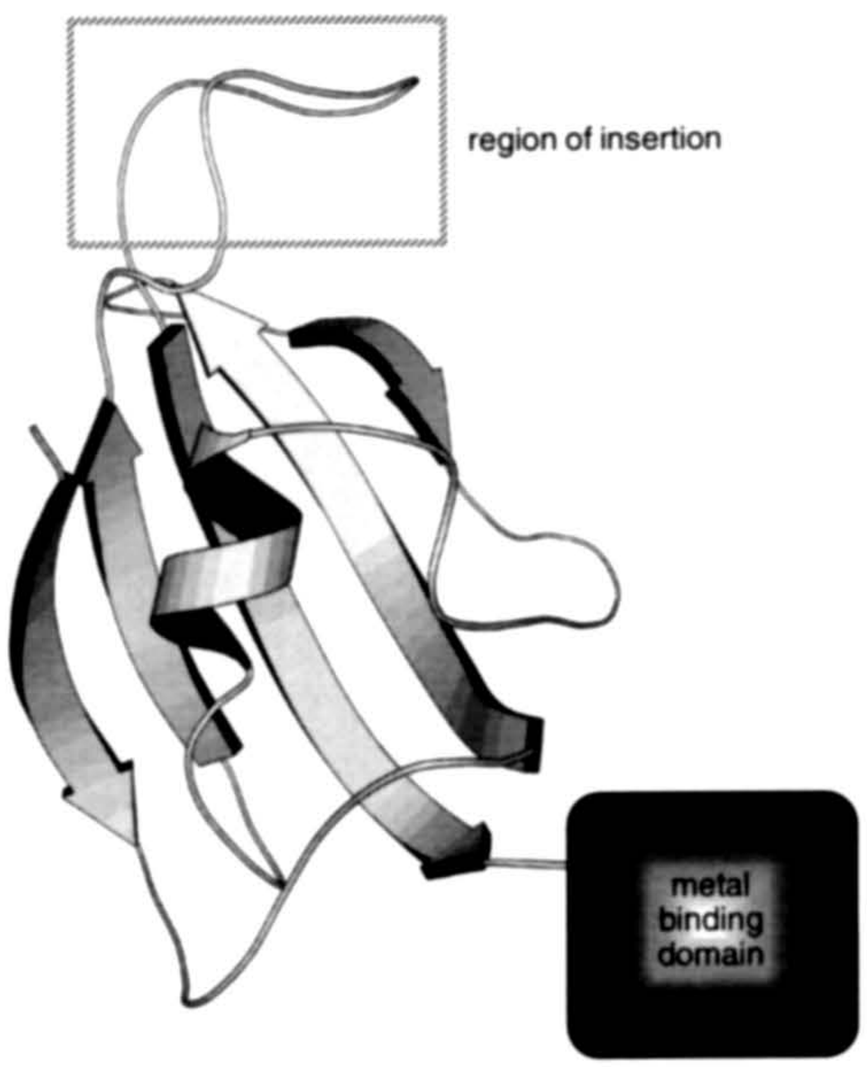

FIG. 5. **Proposed overall structure of WHP.** The structure is based on the human FK506-binding protein structure (van Duyne *et al.*, 1991; Brookhaven protein data base entry 1FKF) and drawn with MOLSCRIPT (Kraulis, 1991). The *box* on *top* shows where the large insertion of WHP and the two homologous translation products of the ORFs, compared with FKBP, would lie in the structure. The independent COOH-terminal metal binding motif is represented by the *box* on the *right*.

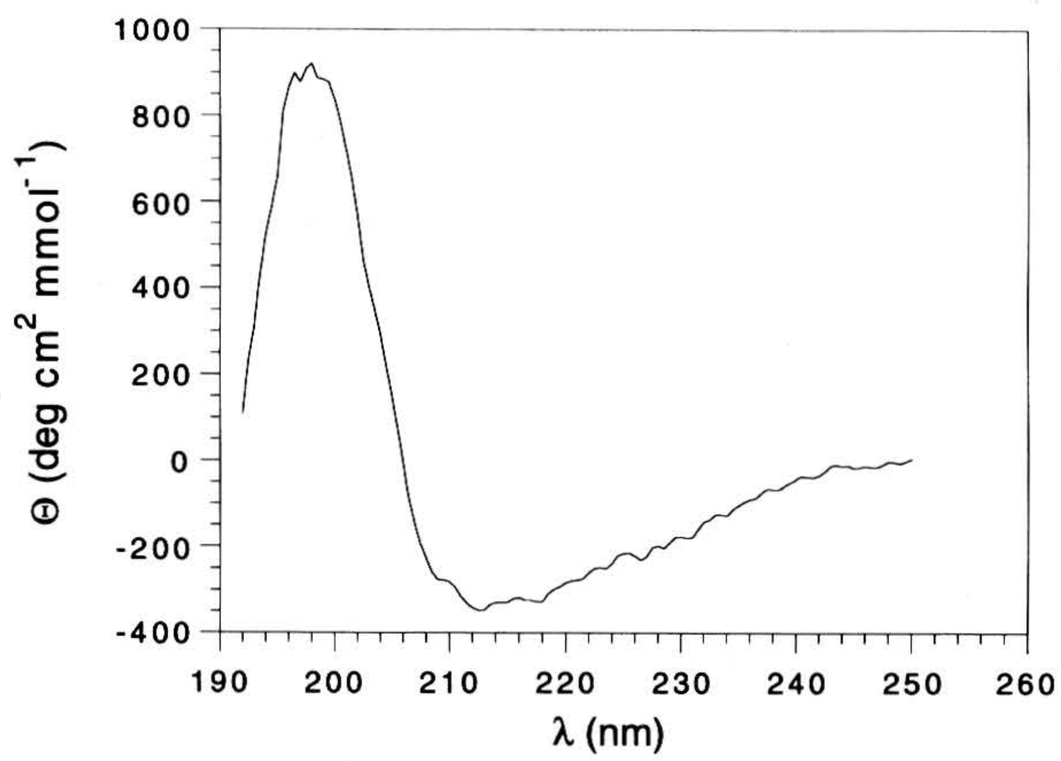

FIG. 6. **Circular dichroism spectrum of WHP.**

transition metal ions has been shown. For Zn^{2+} and Ni^{2+}, which are carried through at least one complete purification procedure, it is probable that they are bound *in vivo*. Judging from the metal binding experiments, Ni^{2+} and Zn^{2+} have the highest binding constant of all metals tested. To conclusively answer the question which metals are bound *in vivo*, the binding constants of different metals have now to be determined.

Usually divalent metal ions are coordinated by a finite small number of chelating amino acids often distant in sequence (Glusker, 1991; Christianson, 1991). In contrast, the COOH-terminal domain of WHP essentially consists of a stretch of potentially metal-binding amino acids: 28 out of 50 amino acids can potentially take part in chelating metal ions. This is reflected by the ability of WHP to bind different metals such as Ni^{2+}, Zn^{2+}, Co^{2+}, Cu^{2+}, and Ca^{2+}, at least to some extent. Being flexible enough to accommodate metal ions with rather different coordination requirements, WHP might not be able to use the bound metal ion as a permanent structural anchor. It rather seems possible that the COOH-terminal part might exist in several energetically similar structures, depending on the nature of the metal ion bound and its preferred coordination. Since the COOH-terminal motif is devoid of hydrophobic amino acids, it is questionable whether the domain has a defined intrinsic fold. In contrast, the binding of a metal ion might only induce a defined structure, making the motif a sensor. The COOH-terminal part of WHP is rich in histidines in close proximity of acidic amino acids, which might shift the pK_a of the histidines to more alkaline values. The actual charge of the histidines, and thereby their chelating ability, may therefore depend on small pH changes in the physiological range. It is suggested that the motif functions as a metal or proton sensor by temporary metal ion binding inducing a specific fold. This in turn might either influence the activity of the NH_2-terminal domain or the interaction with other proteins.

Does WHP fulfill this putative task in an unbiased fashion regarding different divalent metal ions? WHP binds different divalent metal ions, preferentially Ni^{2+} and Zn^{2+}. Whereas Zn^{2+}-binding proteins are ubiquitous, only four classes of Ni-binding enzymes have been described so far: hydrogenases (Hausinger, 1987), CO-dehydrogenases (Conover *et al.*, 1990), methyl-coenzyme M-reductases (Zimmer and Crabtree, 1990) and ureases (Dixon *et al.*, 1975). Of these enzymes, only hydrogenases are known to occur in *E. coli*. The oxidation of molecular hydrogen is nickel-dependent (Evans *et al.*, 1987). Three hydrogenases have been discovered in *E. coli*, and the genes are each arranged in operons with other genes of unknown functions (Menon *et al.*, 1990; Böhm *et al.*, 1990; Przybyla *et al.*, 1992). Another operon contains functions essential for all three hydrogenases (Lutz *et al.*, 1991). The existence of a cluster of metal-binding amino acids (ORF in the hydrogenase operon of *R. leguminosarum* whose NH_2 terminus is similar to the COOH terminus of WHP) in a protein, probably with an accessory role in hydrogenase activity, makes it tempting to speculate that WHP might also be involved in related functions. Since WHP is abundant in the cell and the gene of WHP is not part of an operon, however, as it seems to have a promoter and terminator flanking it, a more general role of this protein is probable.

In summary, it is suggested that the COOH-terminal part of WHP acts as a sensor modulating the function of the NH_2-terminal domain, which may be a substrate-specific rotamase. The COOH terminus may respond to more general changes in physiology by binding Zn^{2+}, pH-dependent or not, or to a more specific situation binding Ni^{2+} *in vivo*.

In addition, WHP was found as the product of a gene required for phage lysis (termed slyD) (Roof *et al.*, 1993). In this case WHP is necessary for the phage gene E-mediated cell lysis to occur.

Even while the cellular function of WHP has not yet been established, it may deserve some further investigation. First, the determination of the structure of the NH_2-terminal domain may test the hypothesis of the FKBP fold and lead to valuable insight how the FKBP structure as a module may be extended by enlarging a loop. Second, the small size of the COOH-terminal part may make it a useful module to confer general divalent metal ion binding ability with a high binding constant to any protein of interest. Third, nuclear and electronic magnetic resonance, as well as extended x-ray absorption fine structure spectroscopy may be used to gain insight into the way various cations are bound. Particularly interesting is the question whether different metals can choose their own coordinating residues from the large offering of this motif. These investigations may be especially useful for metals for which only few natural models are available, such as Ni^{2+}. Finally, it is unresolved whether a motif with essentially no hydrophobic residues has a defined structure at all.

In conclusion, our evidence suggests that WHP is a two-domain protein being composed of one structurally rigid domain with a fold partially similar to FKBP, and one flexible motif being able to bind metal ions, possibly as a means of modulating the function of the NH_2-terminal domain, in response to Zn^{2+}, Ni^{2+}, or pH.

Acknowledgments—First and foremost we thank Dr. Ry Young and his co-workers for generously sharing unpublished data with us. We, furthermore, would like to thank Dr. Josef Kellermann for peptide sequencing, Peter Lindner for help with the CD spectroscopy, Helmut Hartl for carrying out the AAS measurements, Dr. H. Bujard for the plasmid pUHE25-2, and Dr. R. J. P. Williams for helpful discussions.

REFERENCES

Arnold, F. H. (1991) *Bio/Technology* **9,** 151–156
Beveridge, T. J. (1989) in *Metal Ions & Bacteria* (Beveridge, T. J., and Doyle, R. J., eds) pp. 1–30, John Wiley & Sons, New York
Böhm, R., Sauter, M., and Böck, A. (1990) *Mol. Microbiol.* **4,** 231–243
Brendel, V., and Trifonov, E. N. (1984) *Nucleic Acids Res.* **12,** 4411–4427
Brent, R., and Ptashne, M. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4204–4208
Christianson, D. W. (1991) *Adv. Protein Chem.* **42,** 281–355
Conover, R. C., Park, J.-B., Adams, M. W. W., and Johnson, M. K. (1990) *J. Am. Chem. Soc.* **112,** 4562–4564
Dixon, N. E., Gazzola, C., Blakeley, R. L., and Zerner, B. (1975) *J. Am. Chem. Soc.* **97,** 4131–4133
Evans, H. J., Harker, A. R., Papen, H., Russell, S. A., Hanus, F. J., and Zuber, M. (1987) *Annu. Rev. Microbiol.* **41,** 335–361
Fischer, G., Bang, H., and Mech, C. (1984) *Biomed. Biochim. Acta* **43,** 1101–1111
Gebeyehu, G., Rao, P. Y., SooChan, P., Simms, D. A., and Klevan, L. (1987) *Nucleic Acids Res.* **15,** 4513–4534
Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* **182,** 319–326
Glusker, J. P. (1991) *Adv. Protein Chem.* **42,** 1–76
Gribskov, M., Lüthy, R., and Eisenberg, D. (1989) *Methods Enzymol.* **183,** 146–159
Harding, M. W., Handschumacher, R. E., and Speicher, D. W. (1986) *J. Biol. Chem.* **261,** 8547–8555
Harley, C. B., and Reynolds, R. P. (1987) *Nucleic Acids Res.* **15,** 2343–2361
Hausinger, R. P. (1987) *Microbiol. Rev.* **51,** 22–42
Hayano, T., Takahashi, N., Kato, S., Maki, N., and Suzuki, M. (1991) *Biochemistry* **30,** 3041–3048
Hidalgo, E., Palacios, J. M., Murillo, J., and Ruiz-Argüeso, T. (1992) *J. Bacteriol.* **174,** 4130–4139
Higgins, D. G., Bleasby, A. J., and Fuchs, R. (1992) *Comput. Appl. Biosci.* **8,** 189–191
Ishii, S. I., Yokosawa, H., Kumazaki, T., and Nakamura, I. (1983) *Methods Enzymol.* **91,** 378–383
Kainz, P., Schmiedlechner, A., and Strack, H. B. (1992) *Anal. Biochem.* **202,** 46–49
Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
Lindner, P., Guth, B., Wülfing, C., Krebber, K., Steipe, B., Müller, F., and Plückthun, A. (1992) *Methods* **4,** 41–56
Liu, J., and Walsh, C. T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 4028–4032
Liu, J., Albers, M. W., Chen, C., Schreiber, S. L., and Walsh, C. T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 2304–2308
Liu, J., Farmer, Jr., J. D., Lane, W. S., Friedman, J., Weissman, I., and Schreiber, S. L. (1991) *Cell* **66,** 807–815
Lüthy, R., Bowie, J. U., and Eisenberg, D. (1992) *Nature* **356,** 83–85
Lutz, S., Jacobi, A., Schlensog, V., Böhm, R., Sawers, R., and Böck, A. (1991) *Mol. Microbiol.* **5,** 123–135
Menon, N. K., Robbins, J., Peck, H. D., Jr., Chatelus, C. Y., Choi, E.-S., and Przybyla, A. E. (1990) *J. Bacteriol.* **172,** 1969–1977
Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 2444–2448
Perry, A. C. F., Nicolson, I. J., and Saunders, J. R. (1988) *J. Bacteriol.* **170,** 1691–1697
Ponce, M. R., and Micol, J. L. (1992) *Nucleic Acids Res.* **20,** 623
Porath, J. (1987) *Biotechnol. Prog.* **3,** 14–21
Provencher, S. W. (1982) *Comput. Physics Comm.* **27,** 229–242
Przybyla, A. E., Robbins, J., Menon, N., and Peck, H. D., Jr. (1992) *FEMS Microbiol. Rev.* **88,** 109–136
Roof, W. D., Horne, S., Young, K. D., and Young, R. (1993) *J. Biol. Chem.* 269, 2902–2910
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
Scherer, G. F. E., Walkinshaw, M. D., Arnott, S., and Morré, D. J. (1980) *Nucleic Acids Res.* **8,** 3895–3907
Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, C. S., and Sigal, N. H. (1989) *Nature* **341,** 755–757
Silver, J. (1991) in *PCR-A Practical Approach* (McPherson, M. J., Quirke, P., and Taylor, G. R., eds) pp. 137–146, IRL Press, Oxford
Stormo, G. D., Schneider, T. D., and Gold, L. M. (1982) *Nucleic Acids Res.* **10,** 2971–2996
van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1991) *Science* **252,** 839–842
Zimmer, M., and Crabtree, R. H. (1990) *J. Am. Chem. Soc.* **112,** 1062–1066